



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA QUÍMICA  
ENG07053 - TRABALHO DE DIPLOMAÇÃO EM  
ENGENHARIA QUÍMICA



# **Reconciliação de Dados com Discrepância de Modelo Aplicada a uma Coluna de Destilação**

*Autor: Guilherme Bittencourt da Fontoura*

*Orientador: Prof. Dr. Pedro Rafael Bolognese Fernandes*

Porto Alegre, janeiro de 18

Sumário

Agradecimentos	iii
Resumo	iv
Lista de Figuras	v
Lista de Tabelas	vi
Lista de Símbolos	vii
Lista de Abreviaturas e Siglas	viii
1 Introdução	1
2 Revisão Bibliográfica	3
2.1 Medições e Tipos de Erros	3
2.2 Classificação de Variáveis	4
2.3 Reconciliação de Dados	5
2.3.1 Reconciliação de Dados Clássica	5
2.3.2 Reconciliação de Dados Aplicada a Processos de Separação	6
3 Materiais e Métodos	11
3.1 Descrição do Modelo da coluna de Destilação	11
3.2 O Equilíbrio Termodinâmico: Volatilidade Relativa	13
3.3 Reconciliação de Dados com Discrepância no Modelo	13
3.3.1 Solução estacionária da coluna - “Dados Nominais”	14
3.3.2 Geração dos ruídos à solução estacionária	15
3.3.3 Geração da Discrepância no Modelo Termodinâmico	16
3.3.4 Resolvendo o problema da reconciliação de dados	16
4 Resultados	18
4.1 Efeito da Reconciliação de Dados sobre as Vazões	19
4.2 Efeito da Reconciliação de Dados sobre as Composições	22
5 Conclusões e Trabalhos Futuros	27
6 Referências	29

## **Agradecimentos**

Ao meu pequeno, porém gigante, núcleo familiar instalado em Porto Alegre pelo apoio durante toda essa longa etapa chamada engenharia. A minha mãe Kátia pelo exemplo máximo de dedicação, criticidade e retidão. Ao meu irmão Antonio Vicente por ser motivo de muito orgulho e inspiração. Ao meu pai Manoel Antonio e Silva da Fontoura ainda constantemente presente nas nossas memórias e corações.

Ao professor Pedro, meu orientador, pelo exemplo competência, disposição, tranquilidade e bom humor sem os quais esse trabalho seria impossível.

Ao Aor e à Renata, sempre muito presentes e com muito alto astral, colocando sempre a amizade em primeiro lugar.

Aos amigos Shay, Lucas, Henrique, Gatti, Moreschi, Gabriel “Goethe” e Lourenço “Enzão”, irmãos de longa data, pelo companheirismo.

Ao grupo de amigos “desde pequenos” que se formou, companheiros de caminhada das engenharias e de tantos churrascos no “Pupi”, minha gratidão.

A todos que, de alguma forma, se agregaram à minha vida nessa etapa iniciada em 2011.

## **Resumo**

Com o desenvolvimento de instrumentos mais precisos e melhores técnicas de medição é cada vez maior a quantidade de dados disponíveis nos processos industriais. Essa quantidade de dados pode trazer benefícios para a operação e controle dos mesmos, mas não necessariamente está próxima o suficiente da realidade, estando sujeita a erros, sejam eles aleatórios ou sistemáticos. Dados brutos provenientes de medições normalmente não respeitam balanços de massa, de energia e relações termodinâmicas. Os procedimentos de retificação de dados são, então, cada vez mais importantes para o monitoramento e controle de qualquer planta, seja por questões de segurança, de qualidade ou de economia. Nesse sentido, a modelagem e simulação do processo se faz indispensável e é muitas vezes um desafio, uma vez que nenhum modelo descreve perfeitamente o comportamento da planta. Neste presente trabalho é estudado o comportamento da reconciliação de dados aplicada a uma coluna de destilação da mistura BTX (Benzeno, Tolueno e Xileno) em estado estacionário quando há erros no modelo termodinâmico utilizado para prever a separação entre as fases. Chegou-se a resultados satisfatórios, apesar do erro no modelo, com erros máximos de 3,85% na vazão de refluxo e de 0,15% para a composição de benzeno na corrente do destilado. Entretanto, algumas distorções na distribuição das reconciliações em relação à distribuição normal são observadas e aqui discutidas.

## Lista de Figuras

Figura 2.1: Sistemas de separação estudados por Rao e Narasimhan (1996). (a) coluna debutanizadora, (b) planta de extração de suco sintético e (c) circuito de beneficiamento mineral.....	8
Figura 2.2: Esquema da destilação de óleo cru utilizado .....	9
Figura 3.1: Representação do modelo de coluna de destilação utilizado.....	11
Figura 3.2: Esquema da reconciliação de dados com discrepância no modelo termodinâmico .....	14
Figura 4.1: Representação gráfica dos erros absolutos das médias dos valores reconciliados para as variáveis das vazões.....	20
Figura 4.2: <i>normplot</i> para o produto de fundo com $\beta = 0,8$ evidenciando os dois pontos retirados e a proximidade da distribuição normal .....	22
Figura 4.3: Erro na média dos valores reconciliados para as composições do destilado ..	23
Figura 4.4: <i>normplot</i> para $X_{D,xileno}$ e $\beta = 1,3$ com valores zerados e uma distribuição aproximadamente normal para valores não nulos .....	25
Figura 4.5: <i>normplot</i> para $X_{D,benzeno}$ e $\beta = 1,0$ .....	25
Figura 4.6: <i>normplot</i> para $z_{xileno}$ e $\beta = 0,9$ .....	26

## **Lista de Tabelas**

Tabela 3:1: Valores de vazão para estimativa inicial.....	14
Tabela 3:2: Parâmetros especificados para as otimizações via estrutura <i>optimset</i> .....	15
Tabela 3:3: Dados reais da coluna .....	15
Tabela 4:1: Resultados das reconciliações que obtiveram sucesso .....	18
Tabela 4:2: Comparação entre o desvio padrão dos ruídos com o desvio padrão das reconciliações .....	19
Tabela 4:3: Testes da distribuição dos valores reconciliados para as vazões .....	21
Tabela 4:4: Testes da distribuição dos valores reconciliados para as composições .....	24

## **Lista de Símbolos**

$e$  = erro total na medição

$x$  = valor real

$y$  = valor medido

$\varepsilon$  = parcela aleatória do erro

$\delta$  = parcela sistemática do erro

$N_c$  = número de componentes

$N_s$  = número de estágios

$N_f$  = número do estágio de alimentação

$F$  = vazão molar de alimentação

$R$  = vazão molar de refluxo

$D$  = vazão molar de destilado

$B$  = vazão molar de produto de fundo

$V$  = vazão molar de vapor que retorna do refeedor para a coluna

$W_i$  = vazão molar interna de vapor no estágio  $i$

$L_i$  = vazão molar interna de líquido no estágio  $i$

$z_n$  = composição da alimentação para o componente  $n$

$x_{i,n}$  = composição de equilíbrio para a fase líquida do componente  $n$  no estágio  $i$

$y_{i,n}$  = composição de equilíbrio para a fase vapor do componente  $n$  no estágio  $i$

$P$  = pressão

$q$  = qualidade

$\alpha_n$  = volatilidade relativa do componente  $n$  na mistura

$N_{pc}$  = número de períodos das composições

$N_{pf}$  = número de períodos das vazões

$std_f$  = desvio padrão especificado para as vazões

$std_c$  = desvio padrão especificado para as composições

$\beta$  = fator multiplicativo gerador da discrepância planta-modelo

$N_r$  = número de reconciliações que obtiveram sucesso

## **Lista de Abreviaturas e Siglas**

BTX – Benzeno – Tolueno - Xileno

RTD – Retificação de Dados

RD – Reconciliação de Dados

EG – Erros Grosseiros

DEG – Detecção de Erros Grosseiros

MQP – Mínimos Quadrados Ponderados

RRD – Reconciliação Robusta de Dados

EMSO – *Environment for Modelling, Simulation and Optimization*

SARES – Soma Absoluta dos Resíduos

SSE – *Sum of Squared Error*



## 1 Introdução

Para que se obtenha bons resultados em uma planta é fundamental a medição de um número suficiente de variáveis de processo. Segundo Bascur e Linares (2006) um dos maiores desafios para a análise de desempenho operacional é a aquisição, validação e reconciliação de informações de processo. Apesar dos avanços nos sistemas de obtenção de dados, ainda é difícil implementar um sistema de validação de dados para dados de processo, que são muito sensíveis a diferentes fontes de erros, desde erros manuais na obtenção até erros de medidas de sistemas automatizados. Informações validadas e seguras são necessárias para basear qualquer decisão em um negócio, tornando necessário, então, um sistema de reconciliação das informações que unifique e valide os dados, transformando medições brutas em dados mais confiáveis, a fim de auxiliar no controle da planta e na tomada de decisões referentes a ela. A enorme quantidade de dados, a baixa confiabilidade nos dados disponíveis e a falta de consistência dos mesmos são alguns dos problemas mais típicos.

Das técnicas de Retificação de Dados (RTD), a mais comum é a Reconciliação de Dados (RD), que vem sendo utilizada desde a década de 60 (Prata et al., 2010). A técnica consiste em ajustar as medidas de processo visando aumentar a precisão dos dados a partir da propriedade da redundância das variáveis, buscando satisfazer um grupo de restrições, sendo elas equações de balanço de massa, modelos termodinâmicos e restrições impostas enquanto uma função objetivo (medida da diferença entre valores corrigidos e valores observados) é minimizada.

Este trabalho versa sobre a reconciliação de dados aplicada a uma coluna de destilação da mistura benzeno - tolueno - xileno (BTX), na qual o modelo termodinâmico utilizado na reconciliação contém discrepância proposital, quando comparado ao modelo utilizado para se obter os dados nominais, aplicado à volatilidade relativa do Benzeno, composto mais volátil da mistura. Essa discrepância procura simular a situação onde não se conhece perfeitamente a termodinâmica da mistura de interesse, mas sim um modelo aproximado, situação muito comum na indústria. O objetivo é, portanto, observar o comportamento das distribuições dos resultados da reconciliação de dados frente a esse erro a partir de um número suficiente de reconciliações para que se alcance um resultado estatisticamente válido.

O trabalho está estruturado em 5 capítulos. No Capítulo 1 introduz-se o tema a ser abordado. O Capítulo 2 traz informações relevantes da literatura sobre os temas: Medições, Erros e Classificação de Variáveis, assim como noções dos principais temas aqui estudados e aplicações recentes de técnicas de Reconciliação de Dados aplicados à sistemas de separação. O Capítulo 3 apresenta detalhamento da metodologia utilizada, como o modelo da coluna e simplificações consideradas, dados da planta e parâmetros do problema de otimização intrínseco à reconciliação. No Capítulo 4 são apresentados os resultados e, finalmente, no Capítulo 5 as principais conclusões e sugestões de trabalhos futuros na área.

## 2 Revisão Bibliográfica

Neste capítulo é apresentada uma revisão dos conceitos relevantes encontrados na literatura sobre o tema. Inicialmente são abordadas as medições e tipos de erros de medição, seguida de noções de classificação de variáveis e por fim as aplicações da reconciliação de dados.

### 2.1 Medições e Tipos de Erros

As variáveis medidas em um processo, por mais preciso que seja o instrumento de medição, sempre estão associadas a algum erro, sendo somente aproximações ou estimativas do valor real. O erro total na medição ( $e$ ), que é a diferença entre o valor medido ( $y$ ) e o valor real ( $x$ ), pode ser considerado o somatório de duas parcelas: uma aleatória ( $\varepsilon$ ) e uma segunda parcela ( $\delta$ ) que pode conter erros sistemáticos e erros grosseiros, conforme a seguinte definição (BIPM – Bureau International des Poids et Mesures, 2008 apud Zanetti, 2012):

$$e = x - y = \varepsilon + \delta \quad (2.1)$$

O termo aleatório é de natureza probabilística e não pode ter nem a ordem de grandeza nem o sinal previstos, podendo surgir por diferentes motivos: ruídos no ambiente externo, flutuações na rede elétrica, erros na transmissão e transdução dos sinais, filtros, enfim, fatores muitas vezes não controláveis. (Farias, 2009). Erros aleatórios não podem ser completamente eliminados e estão presentes em todas as medições.

Erros sistemáticos são causados por eventos não aleatórios e apresentam magnitude e sinal fixos ou repetíveis, embora geralmente de valor desconhecido. Caso seja conhecido, pode-se usar um fator de correção bias para eliminar seu efeito (BIPM, 2008 apud Zanetti, 2012). Os erros sistemáticos podem ser causados por eventos não aleatórios, como por exemplo o mau funcionamento de instrumentos, mudanças nas condições de medida, calibração falha dos instrumentos, depósito de sólidos, vazamentos ou erros humanos, entre outros (Farias, 2009). Os erros grosseiros podem ser causados por eventos aleatórios e são de maior magnitude, sendo mais facilmente detectáveis por causarem variações pontuais na medição, como por exemplo decorrentes da eventual má leitura da escala por um operador, seja por imperícia ou desatenção.

Em função dos erros de medição, é muito improvável que balanços de massa e energia sejam perfeitamente satisfeitos com os valores obtidos de medidas de processo. Dado que a utilização de valores mais próximos da realidade traz benefícios ao controle, à otimização, à segurança e à eficiência de processos, é essencial a aplicação de algum procedimento de retificação de dados (RTD) para recuperar satisfatoriamente os dados.

### **2.2 Classificação de Variáveis**

As variáveis medidas em um processo industrial podem ser classificadas de acordo com os conceitos de observabilidade e redundância apresentados por Vaclavek e Loucka (1976) e Farias (2009):

- Redundantes – quando a variável continua sendo observável mesmo sem sua medição;
- Não redundantes – quando a variável deixa de ser observável quando não é mais medida.

A classificação de variáveis ajuda a determinar se o problema de reconciliação pode ser resolvido, sendo a informação disponível suficiente, e identificar os conjuntos de variáveis observáveis e não observáveis (Menezes et al., 2014).

O objetivo da classificação, segundo Farias (2009), é eliminar das restrições as variáveis não medidas e aquelas não redundantes (a fim de reduzir o conjunto das restrições). O objetivo da reconciliação de dados pode então ser interpretado como aumentar a precisão dos dados medidos a partir da propriedade da redundância. A diminuição na redundância de uma variável, ainda segundo Farias (2009), faz com que a precisão da reconciliação diminua. A redundância nas medidas é, também, um pré-requisito para Detecção de Erros Grosseiros (DEG), técnica muito utilizada junto com a Reconciliação de Dados, quando há possibilidade também da presença de Erros Grosseiros (sistemáticos) e não só os aleatórios.

## 2.3 Reconciliação de Dados

### 2.3.1 Reconciliação de Dados Clássica

É essencial a aplicação de um procedimento de Retificação de Dados (RTD) para recuperar satisfatoriamente a informação contida em medições de qualquer processo. A RTD é basicamente dividida em 3 etapas:

- Classificação de Variáveis;
- Detecção de Erros Grosseiros;
- Reconciliação de Dados (RD).

O início do uso da RD ocorreu na década de 60, a partir do aumento da capacidade dos computadores. A RD pode ser definida como o ajuste de medidas de processo e parâmetros a fim de satisfazer um grupo de leis de conservação e restrições de processo, enquanto se minimiza uma função objetivo que mede a diferença entre os valores corrigidos e os observados (Prata et al., 2010).

De acordo com Menezes et al. (2014)

O procedimento mais utilizado na RTD é a RD, onde dados medidos são ajustados de maneira estatisticamente coerente pelo estimador (função objetivo) resultante da formulação de máxima verossimilhança sobre a distribuição estatística dos erros de medição assumida, de forma a satisfazer às leis de conservação e demais restrições impostas ao sistema (modelo matemático), obtendo estimativas confiáveis para as variáveis e parâmetros do processo (Prata et al., 2010). Tradicionalmente é assumida distribuição Normal, que resulta no estimador de mínimos quadrados ponderados (MQP).

O problema de otimização é, então, definido tradicionalmente como a minimização do erro quadrático entre as variáveis medidas e as do modelo do processo, conforme Farias (2009):

$$\min_x \left\{ S = (y - x)^T W^{-1} (y - x) \right\} \quad (2.2)$$

sujeito a

$$h(x, p, t) = 0 \quad (2.3)$$

$$g(x, p, t) \leq 0 \quad (2.4)$$

onde  $W$  é uma matriz de ponderação, em geral considerada a matriz de covariância das medições,  $h$  é um conjunto de restrições de igualdade correspondente ao modelo usado e  $g$  é um conjunto de restrições de desigualdade representando os limites operacionais e limites de validade.

O modelo do processo pode ser constituído por equações algébricas no caso da RD estacionária ou de equações algébrico-diferenciais no caso da RD dinâmica. Tais equações representam quaisquer funções que relacionam as variáveis, correspondendo em geral a leis de conservação de massa e de energia. A motivação por trás da utilização das leis de conservação é a baixa probabilidade de existirem falhas na sua estrutura (Crowe, 1996 apud Farias, 2009). O conjunto de restrições pode ser composto de balanços lineares, balanços por componentes ou balanços de energia. Os balanços lineares podem ser, por exemplo, de conservação de massa. Já quando são utilizados balanços por componente apresentam forma bi-linear e quando os balanços são de energia são não-lineares e podem depender do tempo, das variáveis de estado e de parâmetros determinísticos ou aleatórios. Os tipos de restrições, portanto, definem a técnica que será utilizada para a resolução do problema de otimização.

### *2.3.2 Reconciliação de Dados Aplicada a Processos de Separação*

Trabalhos sobre reconciliação de dados encontrados na literatura apresentam aplicações a sistemas multicomponentes. Rao e Narasimhan (1996) salientam o fato de que técnicas de RD são amplamente usadas em diferentes processos químicos e processos de beneficiamento de minerais. Entretanto, conforme os autores, a maioria das aplicações envolve colunas únicas ou uma sequência de unidades de separação. Nesse caso, como o sistema é bi-linear, a função objetivo e as restrições tem suas formas descritas pelas equações 2.5 a 2.7:

$$\min_{\mathbf{F}, \mathbf{x}_j} (\tilde{\mathbf{F}} - \mathbf{F})^T \mathbf{Q}_F^{-1} (\tilde{\mathbf{F}} - \mathbf{F}) + \sum_{j=1}^c (\tilde{\mathbf{x}}_j - \mathbf{x}_j)^T \mathbf{Q}_{x_j}^{-1} (\tilde{\mathbf{x}}_j - \mathbf{x}_j) \quad (2.5)$$

$$\sum_{i=1}^s a_{ki} F_i x_{ij} = 0 \quad k = 1, \dots, n$$

$$j = 1, \dots, c \quad (2.6)$$

$$\sum_{j=1}^c x_{ij} = 1 \quad i = 1, \dots, s \quad (2.7)$$

A função objetivo (2.5) é composta pelas parcelas associadas à vazão ( $F$ ) e à composição ( $x$ ) e as restrições são representadas pelas equações 2.6 e 2.7. As variáveis com sobrescrito “~” representam as medidas de processo,  $n$  é o número de unidades do processo,  $s$  é o número de correntes e  $c$  o número de componentes. O fator de incidência  $a_{ki}$  pode apresentar valores +1, -1 ou 0, dependendo se o termo é entrada, saída ou não incidente, respectivamente, da unidade de processo  $k$ .  $Q_F^{-1}$  e  $Q_{x_j}^{-1}$  são as matrizes de covariância para as vazões e composições, respectivamente. Na formulação acima é assumido implicitamente que os erros nas medidas da vazão não estão correlacionados com os erros nas medidas das composições.

Os autores (Rao e Narasimhan, 1996) comparam três diferentes métodos desenvolvidos na década anterior (Crowe, 1986, Simpson et al., 1988 e Pai e Fisher, 1988) para reconciliação de dados de processos bi-lineares com dados típicos de processos industriais multicomponentes.

- Primeiro o de Crowe (1986), que reduz o problema em dois passos. No primeiro passo as vazões dos componentes correspondentes às composições “não-medidas” são eliminadas usando uma matriz de projeção (Crowe et al., apud Rao e Narasimhan, 1996). No segundo passo vazões totais “não-medidas” são eliminadas usando uma segunda matriz de projeção. Os valores reconciliados das vazões “medidas”, das composições “medidas” da linhas cujas vazões são “não-medidas”, e das vazões por componente “medidas” são obtidas iterativamente. Utilizando os valores reconciliados, as vazões totais “não-medidas” cujas concentrações são “medidas” e “não-medidas” são posteriormente calculadas.

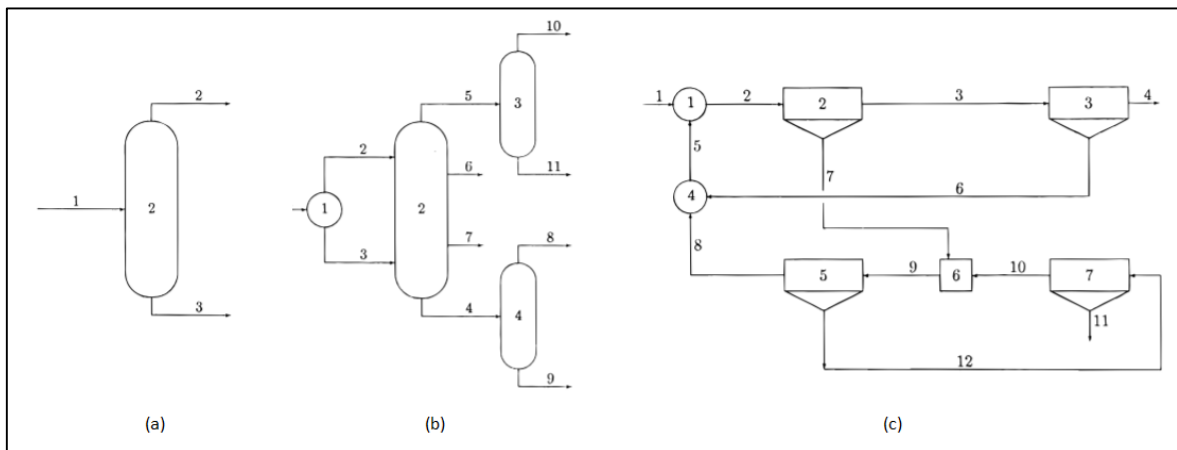
- O segundo, o método de Simpson et al. (1988) reduz-se o problema eliminando todas as restrições e um número igual de variáveis e é definido em três passos. Primeiramente substitui-se as composições na equação 2.5 pelas vazões totais

correspondentes e vazões por componente usando uma aproximação de Taylor de primeira ordem. Então, a partir da definição  $N_{ij} = F_i x_{ij}$ , a equação 2.7 é escrita em função de  $F_i$  e  $x_{ij}$ . Por fim, utilizando os balanços por componente, todas as vazões por componente contidas na função objetivo são escritas em função de um grupo independente de vazões por componente. A função objetivo quadrática sem restrições pode finalmente ser resolvida para obter os valores reconciliados das vazões por componente independentes. As vazões por componente dependentes e as vazões totais dependentes podem, então, ser calculadas.

- O terceiro é uma técnica generalizada não linear descrita por Pai e Fisher (1988).

As reconciliações foram aplicadas aos sistemas de separação mostrados na Figura 2.1.

Figura 2.1: Sistemas de separação estudados por Rao e Narasimhan (1996). (a) coluna debutanizadora, (b) planta de extração de suco sintético e (c) circuito de beneficiamento mineral

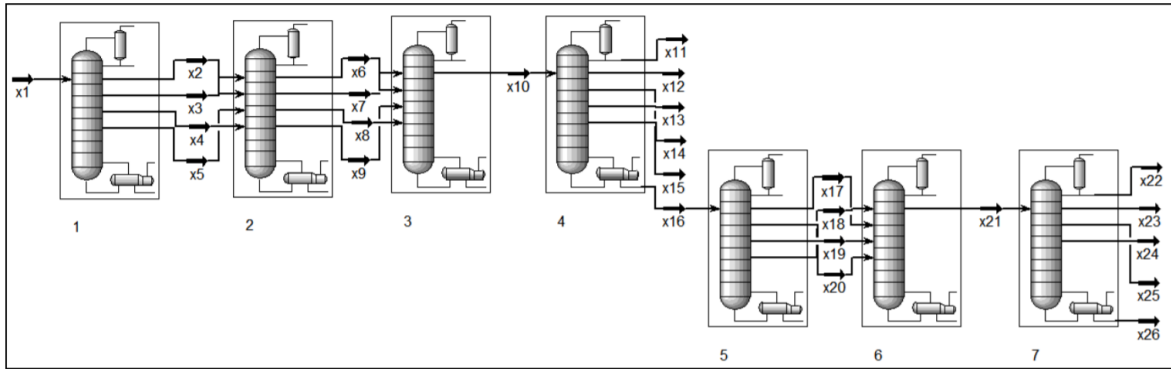


Fonte: Adaptado de Rao e Narasimhan (1996).

Mais recentemente, Menezes et al. (2014) apresentam um estudo sobre reconciliação de dados em estado estacionário em colunas de destilação baseado em trabalhos anteriores na literatura. Um deles é a coluna debutanizadora (Rao e Narasimhan, 1996) e a outra é um conjunto de 7 colunas em série, proposto originalmente por Wang et al. (2004) e ilustrado abaixo, na Figura 2.2.



Figura 2.2: Esquema da destilação de óleo cru utilizado



Fonte: Adaptado de Menezes et al. (2014).

Os autores aplicaram a Reconciliação Robusta de Dados (RRD), na qual é utilizado um estimador robusto (Estimador-M), que são “generalizações de um estimador de máxima verossimilhança” (Prata et al., 2010 apud Menezes et al., 2014). Assume-se que os erros de medição não são correlacionados e o problema de RRD pode ser formulado como

$$\min \sum_i \rho \left( \frac{x_i - z_i}{\sigma_i} \right) = \min \sum_i \rho(\xi_i) \quad (2.8)$$

Sujeito a:

$$h(\underline{x}, \underline{u}) = 0 \quad (2.9)$$

$$g(\underline{x}, \underline{u}) \geq 0 \quad (2.10)$$

onde  $\rho$  representa uma função usada para a formulação do problema de reconciliação (alguns exemplos são dados abaixo, nas equações 2.12 a 2.14),  $\xi_i$  e  $\sigma_i$  são, respectivamente, o resíduo padronizado e o desvio padrão da variável medida.  $z_i$ ,  $\underline{x}$  e  $\underline{u}$  são os vetores das variáveis medidas reconciliadas e não medidas (observáveis) estimadas, respectivamente (Menezes et al., 2014)

Os problemas abordados são resolvidos no pacote computacional EMSO (Soares e Secchi, 2003), que dispõe de rotinas computacionais apropriadas para RD em estado estacionário com restrições lineares ou não lineares. O critério de avaliação escolhido pelos autores foram (Menezes et al., 2014):

- O próprio valor da função objetivo, ou seja, quanto menor, provavelmente melhor terá sido o resultado;

- a Soma Absoluta dos Resíduos (SARES), a saber, a soma absoluta de cada resíduo das restrições - também quanto menor seu valor, provavelmente melhor terá sido o resultado;

- *Sum of Squared Error* (SSE), que nada mais é a soma dos quadrados das diferenças entre os valores reconciliados e os valores exatos ( para x variáveis medidas reconciliadas e u para variáveis não medidas estimadas), conforme demonstrado abaixo (equação 2.11).

$$SSE = \sum (x_{reconciliado} - x_{exato})^2 + \sum (u_{estimado} - u_{exato})^2 \quad (2.11)$$

Os critérios SARES e SSE, segundo os autores, podem ser utilizados para qualquer função objetivo escolhida e, portanto, para fins de comparação entre estimadores e estratégias de RD e DEG.

Menezes et al. (2016) estudaram também a reconciliação robusta em uma coluna com os softwares IISE e SCILAB, apresentando os estimadores-M do tipo *MQP* (não robusto), Fair e Welsch como possíveis escolhas para a função  $\rho$ :

$$\rho_{MQP}(\xi_i) = \frac{\xi_i^2}{2} \quad (2.12)$$

$$\rho_{Fair}(\xi_i, c_F) = c_F^2 \left[ \frac{|\xi_i|}{c_F} - \ln \left( 1 + \frac{|\xi_i|}{c_F} \right) \right] \quad (2.13)$$

$$\rho_{Welsch}(\xi_i, c_W) = \frac{c_W^2}{2} \left[ 1 - \exp \left( -\frac{\xi_i^2}{c_W^2} \right) \right] \quad (2.14)$$

onde  $c_F$  e  $c_W$  são parâmetros de sintonia e são relacionados à eficiência relativa, isto é, quanto mais robusto o estimador, menos eficiente ele é (Albuquerque e Biegler, 1996 apud Menezes et al., 2016).

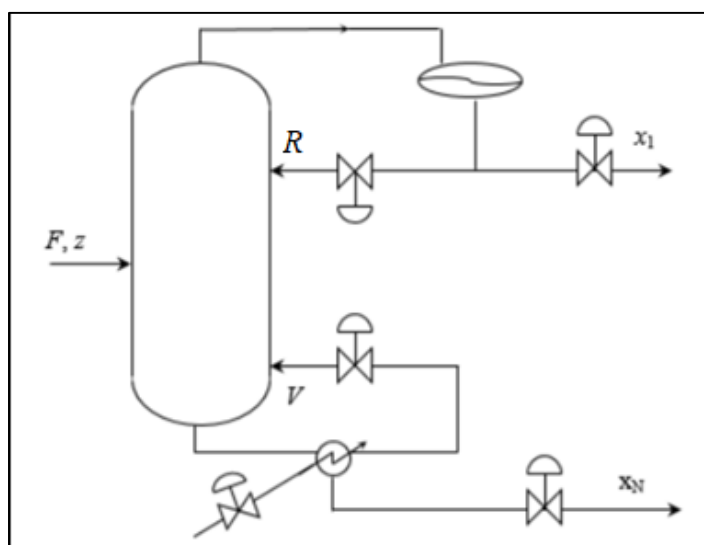
### 3 Materiais e Métodos

Este capítulo apresenta a descrição detalhada do procedimento utilizado para obtenção dos resultados: o modelo da coluna considerada, a metodologia de inserção de discrepância planta-modelo e, por fim, como foi conduzida a reconciliação de dados.

#### 3.1 Descrição do Modelo da coluna de Destilação

A coluna de destilação simulada no software *Matlab* (Versão R2014a) é uma coluna simplificada em estado estacionário conforme hipóteses descritas a seguir. Neste trabalho foi usado o número de componentes ( $N_c$ ) é 3: Benzeno ( $B$ ), Tolueno ( $T$ ) e Xileno ( $X$ ). Como mostra a figura abaixo (Figura 3.1), a vazão molar de alimentação é chamada de  $F$ , a vazão molar de refluxo chamada de  $R$  e a de vapor que retorna do refeedor de  $V$ . As composições da alimentação, do destilado e do produto de fundo são, respectivamente,  $z$ ,  $x_1$  e  $x_N$ . As vazões molares do destilado e do produto de fundo serão chamadas, respectivamente,  $D$  e  $B$ .

Figura 3.1: Representação do modelo de coluna de destilação utilizado



Fonte: Adaptado de Zani (2016).

O modelo utiliza algumas simplificações:

- estágios 100% eficientes;
- equilíbrio termodinâmico entre fases;
- ausência de perdas térmicas;

## 12 Reconciliação de Dados com Discrepância de Modelo Aplicada a uma Coluna de Destilação

- fluxo equimolar, ou seja, vazões internas constantes.

Objetivando variar somente o erro no modelo termodinâmico, que será descrito na seção mais adiante, outros parâmetros são fixados. As frações molares da alimentação para o Benzeno ( $z_{benzeno}$ ), Tolueno ( $z_{tolueno}$ ) e Xileno ( $z_{xileno}$ ) são, respectivamente, 0,61 , 0,25 e 0,14. Assume-se qualidade ( $q$ ) da alimentação igual a 1, ou seja, líquido saturado. A pressão ( $P$ ) é também assumida constante igual a 1 kgf/cm<sup>2</sup>. O número de estágios ( $N_s$ ) é 34, sendo o condensador o primeiro e o refeedor o último, e o estágio de alimentação ( $N_f$ ) é o de número 16.

As vazões molares de líquido ( $L$ ) e de vapor ( $V$ ) nas seções de retificação e de esgotamento seguem o equacionamento abaixo:

$$\begin{aligned} L_i &= R & \text{para } i < N_f \text{ (retificação)} \\ L_i &= R + F q & \text{para } i \geq N_f \text{ (esgotamento)} \end{aligned} \quad (3.1)$$

$$\begin{aligned} W_i &= V + F (1 - q) & \text{para } i \leq N_f \text{ (retificação)} \\ W_i &= V & \text{para } i > N_f \text{ (esgotamento)} \end{aligned}$$

onde  $i$  é o número do estágio,  $R$  é a vazão de refluxo e  $W$  a vazão interna de vapor por estágio. As vazões descendentes de líquido e ascendentes de vapor mostradas na figura acima são aqui representadas por  $L$  e  $V$  e assim chamadas para diferenciar do valor das vazões de refluxo e do refeedor.

Os balanços em cada estágio  $i$ , para um componente  $n$ , já considerando estado estacionário (termos à esquerda da igualdade nas equações, os acúmulos, iguais a zero) são descritas abaixo.

- Para o condensador total:

$$0 = V_2 y_{2,n} - (R + D) x_{1,n} \quad (3.2)$$

- Para a seção de retificação ( $i = 2, \dots, N_c - 1$ ):

$$0 = L_{i-1} x_{i-1,n} + V_{i+1} y_{i+1,n} - L_i x_{i,n} - V_i y_{i,n} \quad (3.3)$$

- Para o estágio de alimentação ( $i = Nf$ )

$$0 = F z_n + L_{i-1} x_{i-1,n} + V_{i+1} y_{i+1,n} - L_i x_{i,n} - V_i y_{i,n} \quad (3.4)$$

- Para a seção de esgotamento ( $i = Nf+1, \dots, Ns-1$ ):

$$0 = L_{i-1} x_{i-1,n} + V_{i+1} y_{i+1,n} - L_i x_{i,n} - V_i y_{i,n} \quad (3.5)$$

Para o refeedor ( $i = Ns$ ):

$$0 = L_{i-1} x_{i-1,n} - B x_{i,n} - V_i y_{i,n} \quad (3.6)$$

onde  $R$  é a vazão molar de refluxo,  $D$  é a vazão molar de destilado e  $B$  é a vazão molar de produto de fundo.

### 3.2 O Equilíbrio Termodinâmico: Volatilidade Relativa

O equilíbrio líquido-vapor foi descrito em cada estágio utilizando volatilidades relativas constantes e é dado por

$$y_n = \frac{\alpha_n x_n}{\sum_{n=1}^{Nc} \alpha_n x_n} \quad (3.7)$$

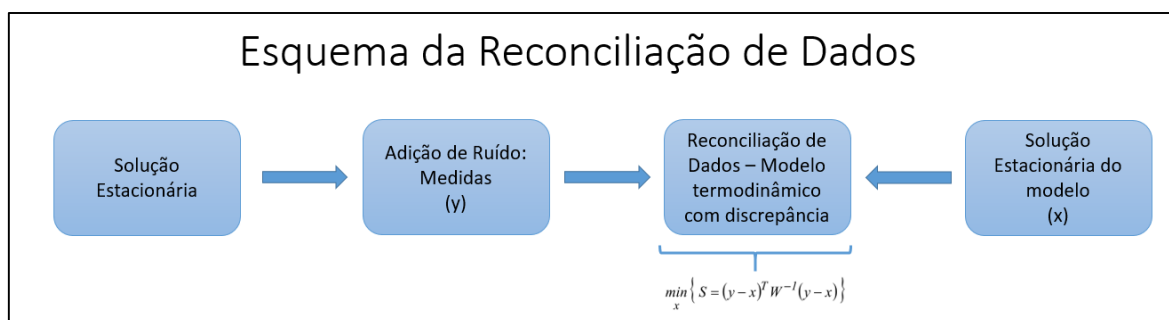
onde  $y_n$  é a fração molar no vapor do componente  $n$ ,  $x_n$  é a fração molar no líquido do componente  $n$  e  $\alpha_n$  é a sua volatilidade relativa. As volatilidades relativas adotadas foram obtidas em uma simulação no software Aspen Plus com dados de entrada nominais e pressão de 1 atm, utilizando o modelo *Gás Ideal / NRTL (Non-Random Two-Liquid)* e calculando-se a média geométrica para as volatilidades obtidas em cada estágio, para ter-se um valor médio e poder utilizá-lo como um parâmetro invariável ao longo da coluna. O resultado para as volatilidades relativas (ao xileno) é  $\alpha_{benzeno} = 5,784165798$ ,  $\alpha_{tolueno} = 2,414653855$  e  $\alpha_{xileno} = 1$ .

### 3.3 Reconciliação de Dados com Discrepância no Modelo

O problema proposto, ilustrado na Figura 3.2, consiste em verificar como se comporta a reconciliação frente a erros de diferentes magnitudes no modelo termodinâmico utilizado para descrever a volatilidade relativa do Benzeno, composto mais volátil em relação aos outros, na mistura BTX. Inicialmente resolvem-se as equações da coluna para uma alimentação pré-fixada e utilizando um modelo da volatilidade relativa sem

discrepâncias propositais. Tendo em mãos esse resultado e considerando-os como “dados reais” do processo, foi adicionado a eles um ruído aleatório com média zero e distribuição normal e dessa distribuição toma-se a média como sendo as “medições” do processo. A partir daí o problema da otimização é conduzido utilizando o modelo corrompido das volatilidades.

Figura 3.2: Esquema da reconciliação de dados com discrepância no modelo termodinâmico



### 3.3.1 Solução estacionária da coluna - “Dados Nominais”

Para obtermos os dados nominais resolvemos as equações de balanço para a coluna em estado estacionário, respeitando a relação termodinâmica anteriormente descrita (equação 3.7) para as volatilidades relativas. A função do *Matlab* chamada para resolver o problema e obter os dados nominais foi a *fsolve* e foram dados como estimativa inicial para a função um vetor com as vazões de estado estacionário descritas na tabela 3.1, com as composições nominais da alimentação e, para todos os outros estágios, com estimativa inicial para todas as composições de 0,333.

Tabela 3.1: Valores de vazão para estimativa inicial

Variável	Valor	Descrição
<b>F</b>	8 kmol/min	Vazão de alimentação
<b>D</b>	5 kmol/min	Vazão molar de destilado
<b>R</b>	12 kmol/min	Vazão de refluxo
<b>V = L + D</b>	17 kmol/min	Vazão de vapor ascendente
<b>B = F - D</b>	3 kmol/min	Vazão de produto de fundo

As opções da otimização adotadas daqui em diante, para limitar as tolerâncias das iterações, são utilizadas via estrutura *optimset* ou *fmincon* (será mencionada mais adiante) e descritos na tabela 3.2:

Tabela 3:2: Parâmetros especificados para as otimizações via estrutura *optimset*

Parâmetro	Descrição	Valor
<b>'TolX'</b>	Tolerância terminal no valor do vetor otimizado 'x0'	1E-06
<b>'TolFun'</b>	Tolerância terminal no valor da função objetivo	1E-06
<b>'TolCon'</b>	Tolerância para violação das restrições	1E-06
<b>'MaxFunEvals'</b>	Número máximo permitido de avaliações da função	1E+05
<b>'MaxIter'</b>	Número máximo permitido de iterações	1E+05

A partir da simulação da coluna em estado estacionário foram obtidos os seguintes valores para os “dados reais”, mostrados na tabela 3.3:

Tabela 3:3: Dados reais da coluna

Variável	Valor	Variável	Valor
<b>D</b>	5,0000	<b>x<sub>D,benzeno</sub></b>	0,9760
<b>R</b>	12,0000	<b>x<sub>D,tolueno</sub></b>	0,0240
<b>B</b>	3,0000	<b>x<sub>D,xileno</sub></b>	0,0000
<b>F</b>	8,0000	<b>x<sub>B,benzeno</sub></b>	0,0000
<b>z<sub>benzeno</sub></b>	0,6100	<b>x<sub>B,tolueno</sub></b>	0,6267
<b>z<sub>tolueno</sub></b>	0,2500	<b>x<sub>B,xileno</sub></b>	0,3733
<b>z<sub>xileno</sub></b>	0,1400	-	-

### 3.3.2 Geração dos ruídos à solução estacionária

Um ruído foi então adicionado a esses dados para simular as oscilações naturais ao processo e obter, então, os valores correspondentes a medidas de processo. O distúrbio é aleatório e segue uma distribuição normal com média zero e desvio padrão especificado. Para cada variável a ser reconciliada foi somado um vetor de valores aleatórios gerado pelo comando *randn*, onde o número de períodos (parâmetro necessário para o comando *randn*) das composições (*Npc*) escolhido foi 10 e das vazões (*Npf*) 50, números considerados suficientemente grandes. O desvio padrão (*std\_f*) escolhido para as vazões foi 10% e para as composições (*std\_c*) 1% em relação ao valor nominal da variável.

Foram obtidos os valores médios (vetor *Média*) e a variância (vetor *Variância*) desses vetores com as distribuições das variáveis. O objetivo de utilizar a média dos valores é evitar a possível geração de erros grosseiros resultantes da geração aleatória dos erros. O vetor *Média* corresponde às medições que serão usadas na reconciliação.

### 3.3.3 Geração da Discrepância no Modelo Termodinâmico

O erro no modelo foi imposto por um fator multiplicativo ( $\beta$ ) aplicado à volatilidade relativa do Benzeno na mistura, como mostra a equação 3.9.

$$\alpha = [ \beta \alpha_{\text{benzeno}} \quad \alpha_{\text{tolueno}} \quad \alpha_{\text{xileno}} ] \quad (3.9)$$

Os valores de  $\beta$  utilizados foram escolhidos entre 0,8 e 1,3, de forma que, a volatilidade relativa do Benzeno ao se aproximar dos valores mais baixos, prevista pelo modelo discrepante, se aproxima bastante da volatilidade do Tolueno na mistura. À medida que o valor de  $\beta$  aumenta o modelo prevê uma volatilidade do benzeno ainda mais distante dos demais componentes.

### 3.3.4 Resolvendo o problema da reconciliação de dados

A reconciliação utiliza o modelo termodinâmico com discrepância e a função do *Matlab* *fmincon* para resolver o problema de otimização.

A estimativa inicial utilizada para a reconciliação ( $s_0$ ) é obtida a partir da solução via *fsolve* do mesmo modelo da coluna, para graus de liberdade especificados a partir dos valores nominais.

A função objetivo a ser reconciliada é da forma

$$\sum_i \left( \frac{(m_i - r_i)^2}{\sigma_i^2} \right) \quad (3.10)$$

onde  $m_i$  são as variáveis medidas (vetor *Média*, vetor das variáveis com ruído),  $r_i$  são os valores reconciliados e  $\sigma_i$  o desvio padrão do vetor das variáveis com ruído da variável  $i$ , onde  $i = 1, \dots, N$  (número de variáveis a serem reconciliadas). As variáveis reconciliadas são obtidas inicialmente resolvendo-se o modelo da coluna



(com a discrepância proposital no modelo termodinâmico descrito na seção anterior) via função *fsolve* do *Matlab* tendo como estimativa inicial os valores do vetor *Média* para as vazões e para a composição da alimentação. As estimativas iniciais das outras composições ao longo da coluna são iguais a 1/3. O resultado obtido é o vetor *s0* utilizado como vetor de partida da otimização que visa minimizar a função objetivo descrita pela equação 3.10.

Na reconciliação, as restrições de igualdade (análogo à equação 2.3) são dadas pelas equações de balanço da coluna e pela relação termodinâmica com o erro proposital descrito na seção anterior. As restrições de desigualdade (equação 2.4) são os limites físicos e operacionais: limites inferiores e superiores de 0 e 1,01 para as composições e para as vazões de 0 e 100 kmol/min.

As variáveis consideradas medidas e que foram reconciliadas são :  $D, R, B, F, Z_{\text{benzeno}}, Z_{\text{tolueno}}, Z_{\text{xileno}}, X_{D,\text{benzeno}}, X_{D,\text{tolueno}}, X_{D,\text{xileno}}, X_{B,\text{benzeno}}, X_{B,\text{tolueno}}, X_{B,\text{xileno}}$ .

As variáveis estimadas são a vazão de vapor  $V$  que retorna do refeedor e as composições de todos os estágios, exceto do destilado e do refeedor.

## 4 Resultados

Este capítulo apresenta uma os resultados obtidos com a reconciliação de dados do modelo de coluna de destilação descrito no capítulo 3, onde um fator de erro ( $\beta$ ) é introduzido no modelo termodinâmico que descreve a volatilidade relativa dos componentes (BTX).

A reconciliação no *Matlab* foi executada 100 vezes para cada valor de  $\beta$ , mantendo o mesmo conjunto nominal de dados, porém com uma diferente realização dos erros de medição simulados. Os resultados aqui mostrados são somente aqueles nos quais a reconciliação obteve sucesso, ou seja, respeitou as restrições e a série de opções de otimização aplicadas via estrutura *optimset* (tabela 3.3). Os valores de  $\beta$  utilizados foram 0,8, 0,9, 1,0, 1,1, 1,2 e 1,3, de forma a não alterar a ordem da volatilidade relativa dos compostos no modelo, uma vez que o mais próximo seria para  $\beta = 0,8$ , onde o produto  $\beta\alpha_{benzeno}$  é aproximadamente 4,6 e  $\alpha_{tolueno} = 2,41$ . Os resultados das reconciliações para cada valor de  $\beta$ , a média dos valores reconciliados, seus desvios padrões, o número de reconciliações com sucesso (*Nr*) e os valores reais das variáveis, para fins de comparação, são apresentados na tabela 4.1.

Tabela 4.1: Resultados das reconciliações que obtiveram sucesso

			$\beta = 0,8$		$\beta = 0,9$		$\beta = 1$		$\beta = 1,1$		$\beta = 1,2$		$\beta = 1,3$	
Variáveis			Nr = 76		Nr = 96		Nr = 99		Nr = 98		Nr = 92		Nr = 94	
Nome	Valor Real	Tipo	Média	Desvio	Média	Desvio	Média	Desvio	Média	Desvio	Média	Desvio	Média	Desvio
'Dest'	5,000	Vazão	5,057	0,056	5,036	0,050	5,022	0,053	5,031	0,058	5,035	0,068	5,040	0,058
'Refluxo'	12,000	Vazão	11,538	0,485	11,629	0,304	11,729	0,304	11,730	0,286	11,662	0,264	11,590	0,268
'ProdFundo'	3,000	Vazão	3,034	0,044	3,030	0,047	3,027	0,056	3,027	0,040	3,030	0,040	3,038	0,047
'Aliment'	8,000	Vazão	8,091	0,088	8,066	0,076	8,049	0,081	8,059	0,076	8,065	0,088	8,078	0,076
'Feed_B'	0,61	Composição	0,6099	0,0012	0,6101	0,0016	0,6102	0,0019	0,6102	0,0014	0,6101	0,0017	0,6100	0,0014
'Feed_T'	0,25	Composição	0,2502	0,0007	0,2500	0,0012	0,2498	0,0014	0,2498	0,0013	0,2498	0,0015	0,2499	0,0013
'Feed_X'	0,14	Composição	0,1400	0,0010	0,1402	0,0009	0,1402	0,0013	0,1400	0,0008	0,1401	0,0010	0,1402	0,0010
'Dest_B'	0,9760	Composição	0,9756	0,0032	0,9770	0,0066	0,9775	0,0074	0,9772	0,0062	0,9771	0,0059	0,9774	0,0072
'Dest_T'	0,0240	Composição	0,0244	0,0032	0,0230	0,0066	0,0225	0,0074	0,0228	0,0062	0,0229	0,0059	0,0226	0,0072
'Dest_X'	8,3E-09	Composição	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
'ProdF_B'	2,8E-07	Composição	0,0004	0,0024	0,0004	0,0015	0,0008	0,0041	0,0003	0,0011	0,0005	0,0031	0,0005	0,0028
'ProdF_T'	0,6267	Composição	0,6264	0,0020	0,6271	0,0031	0,6269	0,0034	0,6270	0,0021	0,6267	0,0025	0,6268	0,0028
'ProdF_X'	0,3733	Composição	0,3734	0,0010	0,3732	0,0020	0,3729	0,0020	0,3728	0,0022	0,3730	0,0017	0,3727	0,0024

À medida que  $\beta$  se afasta do valor 1,0 (valor que corresponde ao caso em que não há discrepância no modelo utilizado para a reconciliação), observou-se uma maior dificuldade na convergência do método, por vezes sendo as iterações interrompidas por violar alguma das restrições, o número máximo de iterações ou o número máximo de avaliações da função, principalmente para valores de  $\beta$  mais baixos, como pode-se ver na tabela 4.1 onde, para  $\beta = 0,8$  o número de reconciliações que convergiram (*Nr*) foi igual a 76, enquanto que para  $\beta = 1,0$  o valor encontrado foi de 99 reconciliações de um total de 100. A razão entre

os desvios padrões das variáveis reconciliadas e os desvios padrões dos valores das medições (vetor do ruído), assim como seu valor, é apresentado na tabela 4.2.

Tabela 4:2: Comparação entre o desvio padrão dos ruídos com o desvio padrão das reconciliações

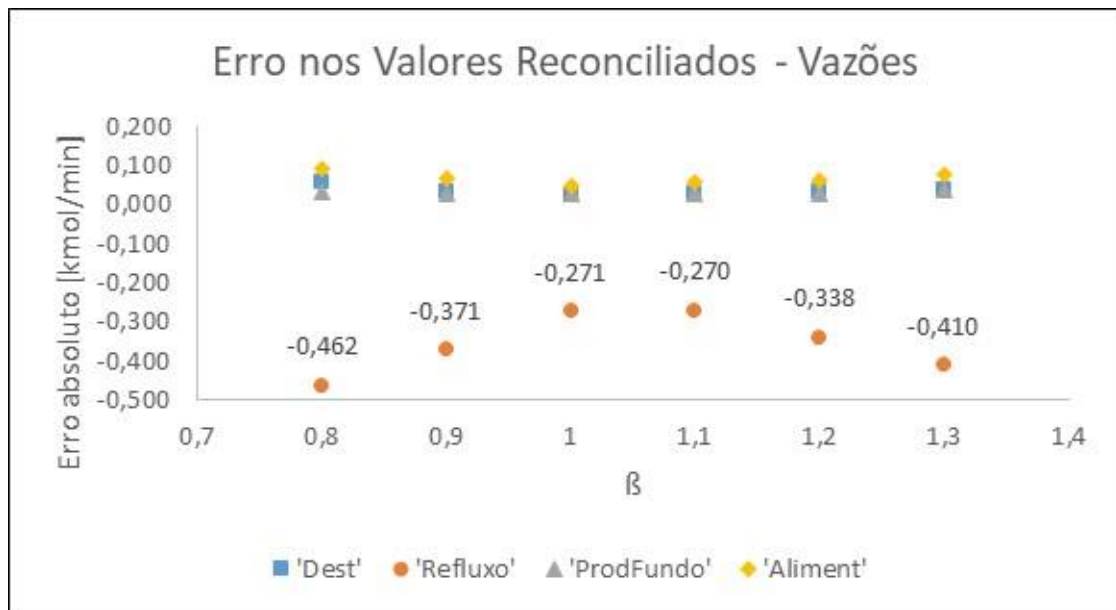
Variáveis			$\beta = 0,8$		$\beta = 0,9$		$\beta = 1$		$\beta = 1,1$		$\beta = 1,2$		$\beta = 1,3$	
			Nr = 76		Nr = 96		Nr = 99		Nr = 98		Nr = 92		Nr = 94	
Nome	Valor Real	Tipo	Desvio Ruído	Desvio Rec. / Ruído	Desvio Ruído	Desvio Rec. / Ruído	Desvio Ruído	Desvio Rec. / Ruído	Desvio Ruído	Desvio Rec. / Ruído	Desvio Ruído	Desvio Rec. / Ruído	Desvio Ruído	Desvio Rec. / Ruído
'Dest'	5,000	Vazão	0,063	0,880	0,070	0,712	0,062	0,862	0,073	0,793	0,078	0,870	0,063	0,920
'Refluxo'	12,000	Vazão	0,158	3,072	0,165	1,849	0,149	2,038	0,162	1,766	0,178	1,486	0,174	1,536
'ProdFundo'	3,000	Vazão	0,033	1,331	0,038	1,259	0,042	1,351	0,041	0,988	0,042	0,957	0,042	1,130
'Aliment'	8,000	Vazão	0,084	1,050	0,124	0,616	0,104	0,778	0,101	0,749	0,113	0,778	0,103	0,737
'Feed_B'	0,61	Composição	0,0020	0,5948	0,0020	0,8058	0,0020	0,9832	0,0019	0,7404	0,0019	0,9098	0,0020	0,7328
'Feed_T'	0,25	Composição	0,0008	0,9242	0,0008	1,4531	0,0008	1,7531	0,0008	1,6691	0,0008	1,9709	0,0008	1,5378
'Feed_X'	0,14	Composição	0,0005	2,1540	0,0004	2,1146	0,0004	3,1779	0,0004	1,8824	0,0004	2,5213	0,0004	2,5724
'Dest_B'	0,9760	Composição	0,0030	1,0701	0,0030	2,1990	0,0031	2,3594	0,0033	1,8680	0,0030	1,9514	0,0031	2,2889
'Dest_T'	0,0240	Composição	0,0001	42,1148	0,0001	100,5382	0,0001	103,9299	0,0001	87,5319	0,0001	80,6870	0,0001	91,7899
'Dest_X'	8,3E-09	Composição	0,0000	131,9462	0,0000	104,9813	0,0000	103,7858	0,0000	77,2716	0,0000	71,6723	0,0000	79,1234
'ProdF_B'	2,8E-07	Composição	0,0000	2532200,1028	0,0000	1757767,4314	0,0000	4805540,4638	0,0000	1333947,6321	0,0000	3313007,9371	0,0000	2785840,6833
'ProdF_T'	0,6267	Composição	0,0021	0,9578	0,0018	1,6971	0,0018	1,8601	0,0015	1,3567	0,0020	1,2716	0,0019	1,4722
'ProdF_X'	0,3733	Composição	0,0011	0,9559	0,0011	1,7707	0,0012	1,7188	0,0012	1,8376	0,0011	1,5448	0,0011	2,0884

#### 4.1 Efeito da Reconciliação de Dados sobre as Vazões

Os desvios padrões observados para as reconciliações não foi em todos os casos menor que o do ruído, como mostrado pelos valores maiores que 1 na tabela 4.2. As maiores discrepâncias foram justamente para o refluxo, variável de vazão com maior desvio padrão nas reconciliações. Já a alimentação e o destilado tiveram valores predominantemente menores que 1, exceto para  $\beta = 0,8$ , onde a razão foi de 1,050.

A partir da Figura 4.1 pode-se observar que, para as variáveis de vazão, desvios em relação ao valor real são mais evidentes para a vazão de refluxo, com uma tendência de aumentar o desvio em relação ao valor real à medida que se intensifica o erro no modelo (fator  $\beta$ ) em ambas as direções. As outras vazões se deslocam no sentido positivo com o aumento da magnitude da discrepância no modelo. O erro, entretanto, é relativamente baixo: com  $\beta$  introduzindo um erro de -20% na volatilidade relativa do benzeno, se observa apenas -0,462 kmol/min de erro para o refluxo  $L$ , que contra a vazão correta de 12 kmol/min, representa apenas um erro de -3,85%. Os erros relativos observados para  $D$ ,  $B$  e  $F$  para o mesmo valor de  $\beta$  são, respectivamente, +1,14%, +1,13% e +1,14%. Ainda vale ressaltar que, respeitando os balanços de massa em estado estacionário, os erros do destilado e do produto de fundo somados resultam no erro da vazão de alimentação.

Figura 4.1: Representação gráfica dos erros absolutos das médias dos valores reconciliados para as variáveis das vazões



A grande variação nas médias das reconciliações para o refluxo pode ser explicado pela correlação que existe entre a razão de refluxo e a pureza dos produtos no topo da coluna. Seu desvio padrão é uma ordem de grandeza mais alto que o das outras vazões (para  $\beta = 0,8$  o desvio padrão do refluxo é 0,485, enquanto que o da alimentação é 0,088), o que ajuda a ilustrar esta dependência do modelo termodinâmico.

Para verificar o efeito do erro de modelo na forma das distribuições das estimativas das variáveis obtidas pela reconciliação, foram realizados os seguintes testes para as distribuições das variáveis reconciliadas:

- Curtose (função *kurtosis* do *Matlab*): indica o quanto mais propensa a distribuição é de apresentar valores distantes da média, em relação à distribuição normal, cuja curtose é 3. Valores maiores indicam uma curva com maior propensão à valores afastados da média e valores menores indicam valores mais concentrados, próximos à media;
- Assimetria (função *skewness* do *Matlab*): Mede a assimetria da curva de distribuição de probabilidades em torno da média. Valores negativos indicam uma maior distribuição à esquerda da média e valores positivos, à direita;
- Testes de Lilliefors (*lillietest*), de Jarque-Bera (*jbttest*) e de Kolmogorov–Smirnov (*kstest*): testam se as distribuições podem ser consideradas do tipo normal, com um nível

de significância de 0,05. A hipótese nula (resultados iguais a zero) é que a distribuição testada seja normal;

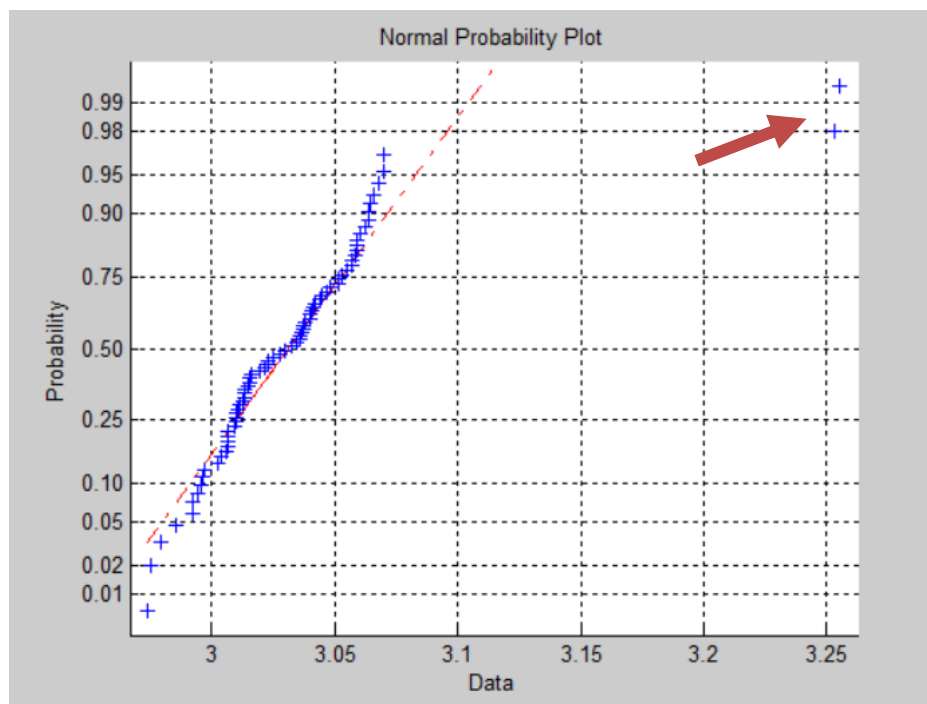
- *Two-Sample t-test (ttest2)*, que testa a hipótese nula de dois vetores virem de amostras aleatórias independentes de distribuição normal com iguais médias, porém com variâncias desconhecidas. No caso da hipótese nula (resultado do teste igual a zero), não se pode descartar a hipótese de que as médias são iguais, no caso, de que não há erro sistemático resultante da reconciliação, com nível de significância de 5%. Esse teste é realizado com dois vetores: o primeiro é o das medidas (dados de planta com ruído) e o segundo é dos resultados das reconciliações com erro no modelo.

Tabela 4:3: Testes da distribuição dos valores reconciliados para as vazões

Variável	Teste	$\beta = 0,8$	$\beta = 0,9$	$\beta = 1,0$	$\beta = 1,1$	$\beta = 1,2$	$\beta = 1,3$
'Dest'	<i>kurtosis</i>	8,33	3,32	2,90	3,45	5,96	4,00
	<i>skewness</i>	1,25	-0,33	-0,27	0,13	-1,24	-0,80
	<i>lillietest</i>	0	0	0	0	1	0
	<i>jbtest</i>	1	0	0	0	1	1
	<i>kstest</i>	0	0	0	0	0	0
	<i>ttest2</i>	1	1	1	1	1	1
'Refluxo'	<i>kurtosis</i>	33,24	6,26	2,72	2,84	2,62	2,67
	<i>skewness</i>	-4,55	-0,78	-0,55	-0,23	-0,12	-0,09
	<i>lillietest</i>	1	0	0	0	0	0
	<i>jbtest</i>	1	1	0	0	0	0
	<i>kstest</i>	1	0	0	0	0	0
	<i>ttest2</i>	1	1	1	1	1	1
'ProdFundo'	<i>kurtosis</i>	16,91	6,02	8,83	4,46	3,86	7,39
	<i>skewness</i>	3,16	1,45	2,04	0,91	0,50	1,59
	<i>lillietest</i>	1	1	1	1	0	1
	<i>jbtest</i>	1	1	1	1	1	1
	<i>kstest</i>	1	0	1	0	0	1
	<i>ttest2</i>	1	1	1	1	1	1
'Aliment'	<i>kurtosis</i>	13,27	2,98	2,49	3,25	2,91	3,81
	<i>skewness</i>	2,05	-0,34	0,07	0,52	-0,40	-0,48
	<i>lillietest</i>	1	0	0	0	1	0
	<i>jbtest</i>	1	0	0	0	0	1
	<i>kstest</i>	0	0	0	0	0	0
	<i>ttest2</i>	1	1	1	1	1	1

A partir das gráficos gerados para o produto de fundo pela função *normplot* do *Matlab* para  $\beta = 0,8$ , como mostrado na figura 4.2, pôde-se observar dois valores resultantes das reconciliações (dos 76 totais) muito discrepantes, que alteram claramente os resultados dos testes da tabela 4.3 para  $\beta = 0,8$ . Os valores da curtose após retirada dos pontos foram recalculados e resultaram em 2,29 para o destilado, 3,04 para o refluxo, 2,04 para o produto de fundo e 2,15 para a alimentação. A figura 4.2 mostra que as distribuições para as vazões não se afastam muito da distribuição normal, descrita pelo tracejado vermelho, exceto os dois pontos mais distantes.

Figura 4.2: *normplot* para o produto de fundo com  $\beta = 0,8$  evidenciando os dois pontos retirados e a proximidade da distribuição normal

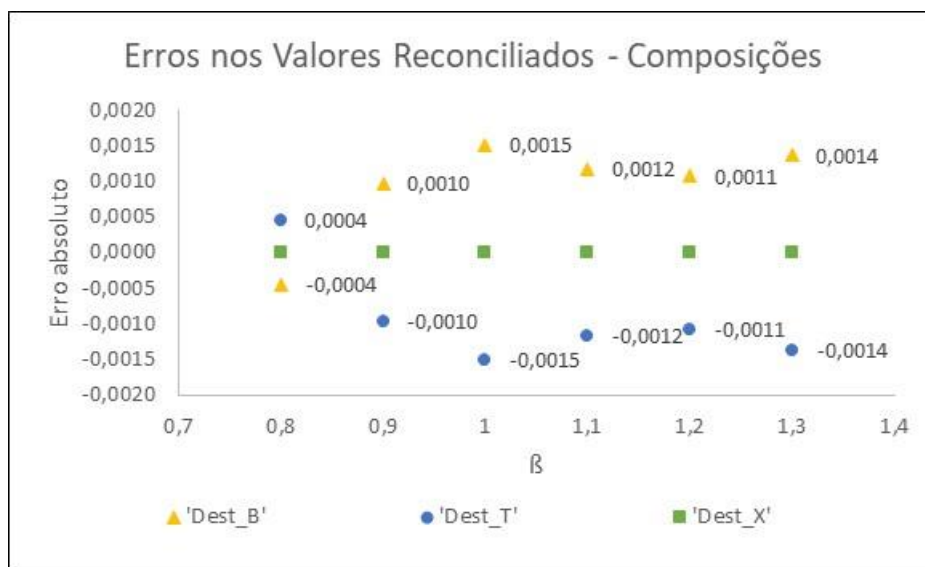


A partir dos testes *lillietest*, *jbttest* e *kstest* da tabela 4.3 pode-se verificar que as distribuições do destilado, do refluxo e da alimentação podem ser consideradas normais, diferentemente do produto de fundo. A maior assimetria do produto de fundo também é notada, onde a cauda direita da distribuição tem maior peso (*skewness* > 0). Ainda, há uma distribuição dos erros entre as variáveis, efeito chamado *smearing*, que pode ser visto pelo erro sistemático presente em todas as vazões, como mostrado pela função *ttest2*, que retornou o valor 1 para todas as variáveis, para todos os valores de  $\beta$ .

#### 4.2 Efeito da Reconciliação de Dados sobre as Composições

Observou-se que, para as variáveis de composição, não há grande prejuízo nas composições reconciliadas. A média dos valores reconciliados é muito próxima dos valores reais e, como mostrado anteriormente na tabela 4.1, os maiores erros absolutos são observados no destilado para as composições do benzeno e do tolueno. O máximo erro verificado foi 0,0015 de erro para as composições do Benzeno e do Tolueno. Esse valor, para uma composição correta de benzeno na corrente do destilado igual a 0,9760, representa apenas 0,15%. O xileno, como esperado, aparece com composição aproximadamente zero. A figura 4.3 traz os erros das médias dos valores reconciliados em função de  $\beta$ .

Figura 4.3: Erro na média dos valores reconciliados para as composições do destilado



As variáveis de composição também tiveram desvios padrão nas reconciliações maior que os desvios do ruído, como mostrado na tabela 4.2. A única composição que ficou com a razão abaixo de 1 para todos os valores de  $\beta$  foi a composição de benzeno na alimentação. As variáveis com maiores da razão dos desvios padrões são as que tem composições próximas ao limite nulo, onde o resultado da reconciliação varia entre valores iguais a zero e valores muito próximos de zero, causando um distúrbio na medida do desvio padrão para essas frações molares nos limites.

Em geral, as variáveis de composição não apresentaram erro sistemático após as reconciliações, como comprovado pelos valores nulos em todos os  $ttest2$  na tabela 4.4, exceto para  $x_{D,xileno}$  e  $x_{B,benzeno}$ .

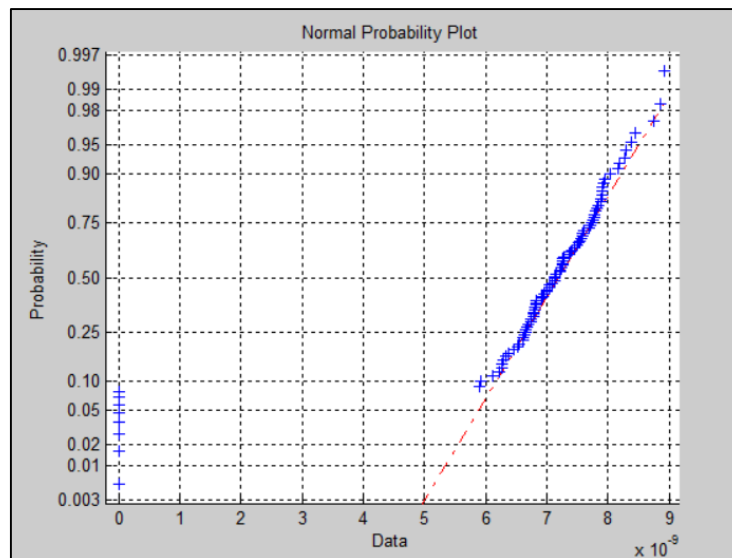
Tabela 4:4: Testes da distribuição dos valores reconciliados para as composições

Variável	Teste	$\beta = 0,8$	$\beta = 0,9$	$\beta = 1,0$	$\beta = 1,1$	$\beta = 1,2$	$\beta = 1,3$
$z_{\text{benzeno}}$	<i>kurtosis</i>	3,18	9,83	6,94	5,87	15,12	3,45
	<i>skewness</i>	0,24	1,94	1,64	1,05	2,95	0,64
	<i>lillietest</i>	0	1	1	0	1	0
	<i>jbtest</i>	0	1	1	1	1	1
	<i>kstest</i>	0	1	1	0	1	0
	<i>ttest2</i>	0	0	0	0	0	0
$z_{\text{tolueno}}$	<i>kurtosis</i>	3,44	6,32	10,51	10,40	26,01	6,60
	<i>skewness</i>	-0,63	-1,51	-2,29	-2,43	-4,21	-1,60
	<i>lillietest</i>	0	1	1	1	1	1
	<i>jbtest</i>	1	1	1	1	1	1
	<i>kstest</i>	0	0	1	1	1	0
	<i>ttest2</i>	0	0	0	0	0	0
$z_{\text{xileno}}$	<i>kurtosis</i>	35,81	11,00	25,63	5,22	41,33	23,85
	<i>skewness</i>	4,87	2,16	3,68	1,08	5,52	3,74
	<i>lillietest</i>	1	1	1	1	1	1
	<i>jbtest</i>	1	1	1	1	1	1
	<i>kstest</i>	1	1	1	0	1	1
	<i>ttest2</i>	0	0	0	0	0	1
$x_{D,\text{benzeno}}$	<i>kurtosis</i>	28,18	10,21	7,82	11,26	12,73	8,72
	<i>skewness</i>	4,10	2,87	2,44	2,95	3,13	2,65
	<i>lillietest</i>	1	1	1	1	1	1
	<i>jbtest</i>	1	1	1	1	1	1
	<i>kstest</i>	1	1	1	1	1	1
	<i>ttest2</i>	0	0	1	0	0	0
$x_{D,\text{tolueno}}$	<i>kurtosis</i>	28,18	10,21	7,82	11,26	12,73	8,72
	<i>skewness</i>	-4,10	-2,87	-2,44	-2,95	-3,13	-2,65
	<i>lillietest</i>	1	1	1	1	1	1
	<i>jbtest</i>	1	1	1	1	1	1
	<i>kstest</i>	1	1	1	1	1	1
	<i>ttest2</i>	0	0	1	0	0	0
$x_{D,\text{xileno}}$	<i>kurtosis</i>	28,61	9,03	7,14	9,95	11,03	8,22
	<i>skewness</i>	3,18	-2,56	-2,20	-2,60	-2,72	-2,49
	<i>lillietest</i>	1	1	1	1	1	1
	<i>jbtest</i>	1	1	1	1	1	1
	<i>kstest</i>	1	1	1	1	1	1
	<i>ttest2</i>	1	1	0	1	1	1
$x_{B,\text{benzeno}}$	<i>kurtosis</i>	50,94	17,78	67,53	20,78	73,44	40,32
	<i>skewness</i>	6,88	3,96	7,65	4,34	8,20	6,12
	<i>lillietest</i>	1	1	1	1	1	1
	<i>jbtest</i>	1	1	1	1	1	1
	<i>kstest</i>	1	1	1	1	1	1
	<i>ttest2</i>	0	1	1	1	0	0
$x_{D,\text{tolueno}}$	<i>kurtosis</i>	3,43	9,58	9,44	5,28	6,29	6,45
	<i>skewness</i>	0,08	2,21	2,27	0,92	1,12	1,62
	<i>lillietest</i>	0	1	1	0	0	1
	<i>jbtest</i>	0	1	1	1	1	1
	<i>kstest</i>	0	1	1	0	0	1
	<i>ttest2</i>	0	0	0	0	0	0
$x_{D,\text{xileno}}$	<i>kurtosis</i>	3,24	9,77	8,36	11,61	13,76	10,70
	<i>skewness</i>	-0,30	-2,19	-2,02	-2,53	-2,69	-2,49
	<i>lillietest</i>	0	1	1	1	1	1
	<i>jbtest</i>	0	1	1	1	1	1
	<i>kstest</i>	0	1	1	1	1	1
	<i>ttest2</i>	0	0	0	0	0	1

O problema com essas variáveis ( $x_{D,\text{xileno}}$  e  $x_{B,\text{benzeno}}$ ) é a proximidade do valor zero e, durante a reconciliação, muitas vezes o valor final não é nulo, mas muito próximo de zero, resultando em uma distribuição assimétrica, aparentando um erro sistemático na função *ttest2*, mas que, na verdade, não pode ser assim considerado. Um exemplo desse problema é demonstrado na figura 4.4, um *normplot* para  $x_{D,\text{xileno}}$  com  $\beta = 1,3$  (observar que os valores não nulos são da ordem de  $10^{-9}$ ):



Figura 4.4: *normplot* para  $X_{D,xileno}$  e  $\beta = 1,3$  com valores zerados e uma distribuição aproximadamente normal para valores não nulos



A tabela 4.4 também mostra os testes *kurtosis*, *skewness*, *lillietest*, *jbtest*, *kstest* e *ttest2* para as variáveis de composição na alimentação ( $z$ ), no destilado ( $x_D$ ) e no produto de fundo ( $x_B$ ). Em geral, os testes não indicaram proximidade da distribuição das composições com a distribuição normal, mas, tirando valores limites (0 ou 1), a distribuição não é muito distante da normalidade para as variáveis mais próximas dos valores limites, como mostrado na figura 4.5 a distribuição dos valores reconciliados para a composição de benzeno no destilado.

Figura 4.5: *normplot* para  $X_{D,benzeno}$  e  $\beta = 1,0$

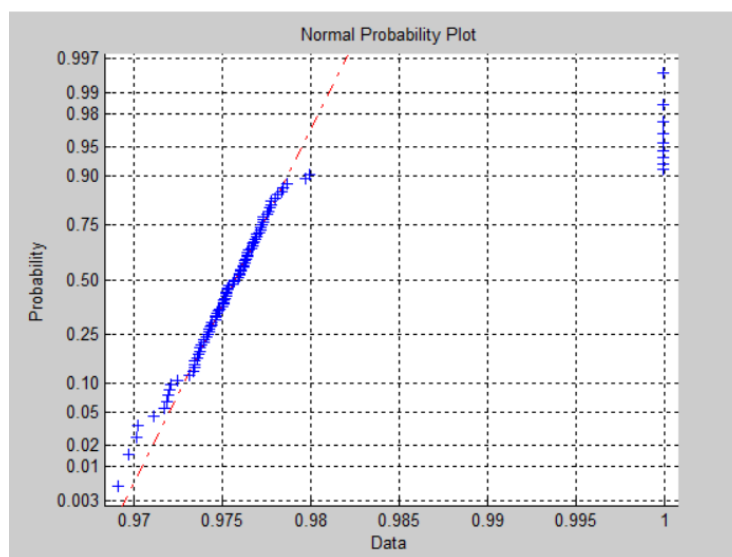
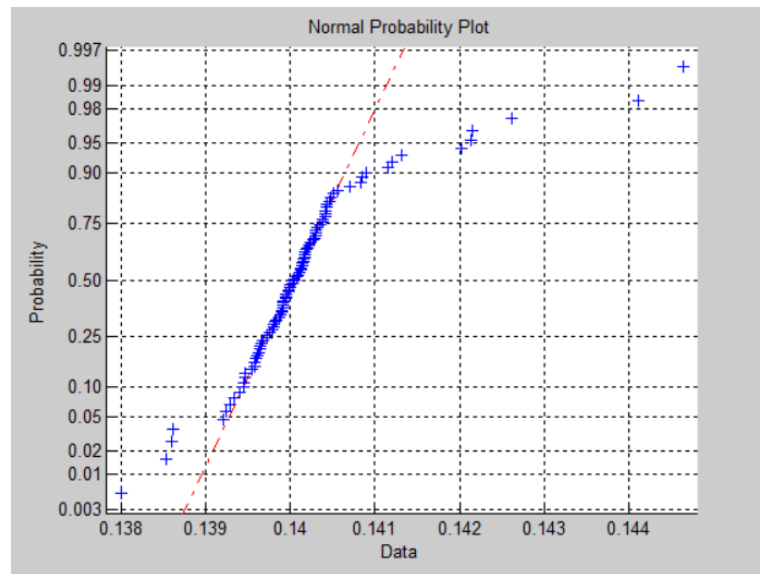


Figura 4.6: *normplot* para  $z_{xileno}$  e  $\beta = 0,9$ 

Aqui na Figura 4.6 os valores apresentam uma propensão a ter sua média deslocada para a direita, gerado pelos valores mais afastados da média, apresentando uma distribuição mais assimétrica, quantificada por maiores valores de assimetria (*skewness*) da tabela 4.4 para as composições, quando comparadas com as vazões (tabela 4.2).

## 5 Conclusões e Trabalhos Futuros

Foi aplicado um range de erro de -20% a +30% no valor da volatilidade relativa do benzeno (composto mais volátil) na mistura BTX, não sendo suficiente para haver alteração da ordem de volatilidade entre os compostos causada pela discrepância do modelo termodinâmico.

Para as variáveis relativas às vazões, as distribuições resultantes foram próximas à distribuição normal, mas há um pequeno erro sistemático no resultado causado pela reconciliação e pelo efeito conhecido como *smearing*, onde há distribuição do erro entre todas as variáveis. As médias das vazões da alimentação, do produto de fundo e do destilado ficaram pouco acima do valor real, sem variar muito com o aumento de  $\beta$ , enquanto que a vazão de refluxo  $R$  apresentou maior variação no sentido negativo com o aumento do erro no modelo (aumento ou diminuição de  $\beta$ ), o que pode ser explicado pela correlação que existe entre a volatilidade relativa e a pureza dos produtos de topo, relacionada à razão de refluxo. Seus valores de curtose são baixos (menores que 3, exceto para o produto de fundo), indicando uma distribuição dos valores reconciliados mais próxima da média, com caudas mais finas.

Para as composições, as reconciliações não geraram erros sistemáticos, como mostrado pelos valores de  $ttest2$  nulos. As distribuições se distanciam da distribuição normal principalmente pelas caudas maiores em uma das direções. As variáveis  $Z_{benzeno}$ ,  $Z_{xileno}$ ,  $X_{D,benzeno}$ ,  $X_{B,benzeno}$  e  $X_{B,tolueno}$  apresentaram caudas da distribuição maiores à direita ( $skewness > 0$ ), enquanto que  $Z_{tolueno}$ ,  $X_{D,tolueno}$ ,  $X_{D,xileno}$  e  $X_{B,xileno}$  apresentaram caudas à esquerda ( $skewness < 0$ ). Os valores de curtose para as composições são maiores que os das vazões, indicando que suas distribuições possuem caudas mais longas, ou seja, maior número de valores mais afastados da média.

As reconciliações foram em geral bem sucedidas, não apresentando grandes erros com pequenas-médias discrepâncias no modelo, exceto para a vazão de refluxo, variável que mais se altera frente à variação do modelo termodinâmico, pois é diretamente relacionada à pureza do topo da coluna.

Trabalhos futuros podem ser feitos no sentido de aproximar a volatilidade relativa de dois compostos ainda mais e verificar o comportamento das distribuições das reconciliações. Também poderia-se variar, ao invés do componente mais volátil, algum dos

outros ou então dois componentes simultaneamente. Pode-se, ainda, verificar se há diferença entre o método clássico utilizado ou a reconciliação robusta de dados, neste caso considerando o erro de modelo como uma componente sistemática do erro.

## 6 Referências

ALBUQUERQUE, J. S., BIEGLER, L. T. **Data reconciliation and gross-error detection for dynamic systems.** AIChE Journal, v. 42, pp. 2841-2856, 1996.

BASCUR, O. A., LINARES, R. **Grade recovery optimization using data unification and real time gross error detection,** Minerals Engineering, v. 19 pp. 696-702, 2006.

BUREAU INTERNATIONAL DES POIDS ET MESURES - BIPM. **Guide to the expression of uncertainty in measurements.** Disponível em <<https://www.bipm.org/en/publications/guides/gum.html>>. Acesso em: 07 de jan. de 2018.

CROWE, C. M. **Reconciliation of process flow rates by matrix projection, part II: the nonlinear case.** AIChE Journal, v. 32, 616-623, 1986.

CROWE, C. M. **Data reconciliation – progress and challenges.** Journal of Process Control, v. 6 pp. 89-98, 1996.

FARIAS, A. C. **Avaliação de estratégias para reconciliação de dados e detecção de erros grosseiros.** Dissertação (Mestrado em Engenharia Química). Porto Alegre: Universidade Federal do Rio Grande do Sul, 2009.

MENEZES, D. Q. F., SARRUF, I., PEIXOTO, F. C. e PRATA, D. M. **Reconciliação de dados em colunas de destilação utilizando o simulador EMSO.** XX COBEQ, 2014.

MENEZES, D. Q. F., SOARES, R., PRATA, D. M., PEIXOTO, F. C., SECCHI, A. R. E PINTO, J. C. C. S. **Reconciliação robusta de dados em coluna de destilação utilizando os softwares ISSE e SCILAB.** XXI COBEQ, 2016.

NARASIMHAN, S., JORDACHE, C. **Data reconciliation and gross error detection: an intelligent use of process data.** Houston: Gulf Publishing Company, 1999.

PAI, C. C. D., FISHER, G. D. **Application of Broyden's method to reconciliation of nonlinearly constrained data.** AIChE Journal, v. 34, pp. 873-876, 1988.

PRATA, D. M.; SCHWAAB, M.; LIMA, E. L.; PINTO, J. C. **Simultaneous Robust Data Reconciliation and Gross Error Detection through Particle Swarm Optimization for an Industrial Polypropylene Reactor**. Chem. Eng. Sci., v. 65, p. 4943-4954, 2010.

RAO, R. R., NARASIMHAN, S. **Comparison of techniques for data reconciliation of multicomponent processes**. Ind. Eng. Chem. Res., v. 35, pp. 1362-1368, 1996.

SIMPSON, D. E., Everett, M. G., Voller, V. R., **An efficient algorithm for mineral processing data adjustment**. Int. J. Miner. Process, v. 31, pp. 73-96, 1991.

SOARES, R. P., SECCHI, A. R. **EMSO: Environment for modelling, simulation and optimization**. Comput. Aided Chem. Eng., v.14, pp. 947-952, 2003.

VACLAVEK, V., LOUCKA, M. **Selection of measurements necessary to achieve multicomponent mass balances in chemical plant**. Chemical Engineering Science, v. 31, pp. 1109-1205, 1976.

ZANETTI, J. P. C. **Detecção de erros grosseiros e reconciliação de dados numa central petroquímica**. Trabalho de Diplomação em Engenharia Química. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2012.

ZANI, J. B. **Reconciliação de dados em colunas de destilação aplicada à estimação da composição da alimentação**. Trabalho de Diplomação em Engenharia Química. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2016.

WANG, F.; JIA, X.; ZHENG, D.; YUE, J. **An improved MT-NT method for gross error detection and data reconciliation**. Comput. Chem. Eng., v. 28, pp. 2189-2192, 2004.